

# Toward Resolving Deep Neoaves Phylogeny: Data, Signal Enhancement, and Priors

Renae C. Pratt,\* Gillian C. Gibb,\* Mary Morgan-Richards,\* Matthew J. Phillips,† Michael D. Hendy,\* and David Penny\*

\*Allan Wilson Centre for Molecular Ecology and Evolution, Massey University, Palmerston North, New Zealand; and †Centre for Macroevolution and Macroecology, School of Botany and Zoology, Australian National University, Canberra ACT, Australia

We report three developments toward resolving the challenge of the apparent basal polytomy of neoavian birds. First, we describe improved conditional down-weighting techniques to reduce noise relative to signal for deeper divergences and find increased agreement between data sets. Second, we present formulae for calculating the probabilities of finding predefined groupings in the optimal tree. Finally, we report a significant increase in data: nine new mitochondrial (mt) genomes (the dollarbird, New Zealand kingfisher, great potoo, Australian owl-nightjar, white-tailed trogon, barn owl, a roadrunner [a ground cuckoo], New Zealand long-tailed cuckoo, and the peach-faced lovebird) and together they provide data for each of the six main groups of Neoaves proposed by Cracraft J (2001). We use his six main groups of modern birds as priors for evaluation of results. These include passerines, cuckoos, parrots, and three other groups termed “WoodKing” (woodpeckers/rollers/kingfishers), “SCA” (owls/potoos/owlet-nightjars/hummingbirds/swifts), and “Conglomerati.” In general, the support is highly significant with just two exceptions, the owls move from the “SCA” group to the raptors, particularly accipitrids (buzzards/eagles) and the osprey, and the shorebirds may be an independent group from the rest of the “Conglomerati.” Molecular dating mt genomes support a major diversification of at least 12 neoavian lineages in the Late Cretaceous. Our results form a basis for further testing with both nuclear-coding sequences and rare genomic changes.

## Introduction

Perhaps, the greatest current challenge of avian systematics for molecular evolutionists and systematists alike is the resolution of the polytomy at the base of the Neoaves. The basic paleognath (tinamous and ratites)—neognath division (all other modern birds) is supported by studies of morphology (Cracraft and Clarke 2001), nuclear-coding DNA (Groth and Barrowclough 1999; García-Moreno and Mindell 2000), and mitochondrial (mt) genomes (Sorenson et al. 2003; Harrison et al. 2004; Slack et al. 2007). Within the Neognathae, the Galloanserae (chickens, ducks, and their relatives) represent the earliest divergence, leaving the large majority (all remaining orders) of birds in the Neoaves. Again, coding regions of both mt genomes and nuclear DNA, together with morphological data, agree with the Galloanserae division. However, resolving the relationships within Neoaves is still elusive.

Thus, resolution of the basal Neoavian polytomy could be seen as the “last frontier” for resolving deep-level systematics among modern birds. There are a range of views best illustrated by the two ends of a spectrum—first, the theory that the basal polytomy is due to an “explosive radiation” after the Cretaceous–Paleogene (K–Pg, formerly K–T) boundary. That is, birds and mammals “inherited the earth” only after the demise of the dinosaurs and pterosaurs (Feduccia 2003; Chubb 2004; Poe and Chubb 2004; Ericson et al. 2006). The other end of the spectrum are hypotheses that basal avian lineages were diversifying in an “adaptive radiation” long before the asteroid impact that marks the K–Pg boundary (Cooper and Penny 1997; Cracraft 2001; van Tuinen and Hedges 2001; Penny and Phillips 2004; Pereira and Baker 2006; van Tuinen et al.

2006; Brown et al. 2007, 2008). This latter approach represents mainstream evolutionary theory in that it attempts to explain the past by reference to known mechanisms—to “causes now in operation” (Penny and Phillips 2004).

Poe and Chubb (2004) suggested that the large polytomy at the base of Neoaves represents a rapid radiation that “might be considered essentially simultaneous.” If a lack of resolution is not caused by truly short times between divergences, then ultimately relationships should be resolvable (Whitfield and Lockhart 2007). We have already commented (Gibb et al. 2007) that an explosive radiation implies both short divergence times between avian orders and also that the ecological and morphological differences that identify the crown groups of orders within Neoaves must have occurred over the same short timescale. It would scarcely be an explosive radiation if the lineages diverged quickly, but it then took tens of millions of years for genetic changes to occur leading to the ecological and morphological characters that distinguish crown group Neoavian orders today. Apart from being real, short branch lengths in phylogenies can result for a number of reasons. For example, different characters or data sources that provide support for conflicting trees (rather than from the absence of support) can result in short branches; under these conditions, even standard maximum likelihood (ML) can seriously underestimate branch lengths (Penny et al. 2008). Additionally, use of inappropriate genes or analysis methods can return short branches. To improve divergence time estimates, Brown et al. (2008) recommend longer sequences (they used 4,594 bp of mtDNA) and the use of multiple independent nuclear loci. In addition to incorporating longer sequences, improved analytic methods such as networks allows visualization of conflict between essentially equally good phylogenies (Holland et al. 2004).

A greater understanding of the evolutionary history of Neoaves is still needed. We have found Cracraft’s (2001) six groupings within Neoaves to be useful as “informal priors” for recent studies on passerines and for the group termed “Conglomerati” (e.g., Gibb et al. 2007). With

Key words: Neoaves, mitochondrial genomes, site-stripping, down-weighting, hypothesis testing, avian evolution.

E-mail: r.c.pratt@massey.ac.nz.

*Mol. Biol. Evol.* 26(2):313–326. 2009

doi:10.1093/molbev/msn248

Advance Access publication November 3, 2008

additional data presented here, we have representatives (and test) his six prior groupings within Neoaves:

- (i) Passeriformes (passerines, perching birds);
- (ii) Psittaciformes (parrots);
- (iii) Cuculiformes (cuckoos);
- (iv) Coliiformes, Coraciiformes, Trogoniformes, and Piciformes (mousebirds [not included in the present study], rollers/bee-eaters/kingfishers, trogons, woodpeckers/toucan, and jacamars). Informally, and for ease of communication, we call this the “WoodKing” grouping;
- (v) Strigiformes, Caprimulgiformes, Apodiformes, and Musophagiformes (owls, nightjars, potoos, owl-nightjars, hummingbirds, swifts, and turacos). We have sampled three of the four groups, so we will refer to these as “SCA” (or “CA” when omitting the owls [Strigiformes]); and
- (vi) a diverse group dubbed the “Conglomerati” (Slack et al. 2007) that includes three main subgroups: 1) Falconiformes, Pelecaniformes, Ciconiiformes, Procellariiformes, Gaviiformes, Sphenisciformes, and Podicipediformes (raptors, pelicans and relatives, storks, seabirds, loons, penguins, and grebes); 2) Charadriiformes, Columbiformes, and Pteroclididae (shorebirds, pigeons, and sandgrouse); and 3) Gruidae, Rallidae, Otidae, Eurypygidae, and Turnices (cranes, rails, bustards, kagu, sun bitterns, and button quails).

In recent papers (Slack et al. 2006, Slack et al. 2007), we concentrated on the relationships within (i) and (vi) but have recently extended this to include members of groups (iv) and (v) (Gibb et al. 2007; Morgan-Richards et al. 2008). Because we previously only had a single representative for the orders Psittaciformes (parrots) and Strigiformes (owls) (Harrison et al. 2004), we omitted them from recent analyses because they are long branches that are known to be problematic in phylogeny generally (Hendy and Penny 1989), including birds (Harrison et al. 2004).

At this stage, we are particularly concerned as to whether the members within each of the above six groups come together—this will help evaluate whether the deep Neoavian lineages are resolvable. As such, we are not especially concerned if some of Cracraft’s six groups have paraphyletic lineages within them—that is, a taxonomic question, not a question about the resolvability of the deepest lineages. For example, we are interested in whether the combined Coliiformes/Coraciiformes/Piciformes (“WoodKing”) form a natural group within Neoaves—even if one of the subgroups turns out to be paraphyletic within this “WoodKing” grouping.

After this work was submitted, Hackett et al. (2008) published what is probably the most comprehensive report on bird evolution since Cracraft (2001). Although the scope of data is impressive, it is largely based on noncoding intron sequences. Nuclear intron data, in combination with current mt and fossil data, have the potential to be extremely useful as long as we can be confident in alignments that span the phylogenetic depth of the avian clade. Some authors have suggested that intron sequences are not appropriate for deeper divergences (Shapiro and Dumbacher 2001) due

to alignment ambiguities resulting from multiple insertions and deletions (for general comments on alignments see Löytynoja and Goldman 2008). For example, Morgan-Richards et al. (2008) showed that the alignment of  $\beta$ -fibrinogen intron 7 (which supports the controversial metaves–coronaves split) has no constant sites across the wide taxon sampling required for determining deep avian divergences. Introns are potentially well suited to resolving rapid radiations as they evolve fast enough to accumulate changes during this time (of divergence), while being slow enough to not become random and therefore lose signal (Matthee et al. 2007).

In general, the main avian orders found in both Hackett et al. (2008) and Cracraft (2001) are the same (e.g., Passeriformes, Psittaciformes, and Cuculiformes etc.). However, the relationships among the orders are different, with the deep branches of Neoaves receiving low support (<80% bootstrap support) in Hackett et al. (2008). One difference is their “land birds.” Within land birds, Hackett et al. (2008) found Passeriformes (passerines) sister to Psittaciformes (parrots) and suggested a sister relationship between these and Falconidae. These groupings only have support when the intron data are included. Ericson et al. (2006) also inferred this relationship but only when all genes were combined, including  $\beta$ -fibrinogen intron 7 (see their Supplementary Material figs. ESM-1–8). If correct, this placement would be very interesting; however, intron alignment and/or long-branch attraction may be a factor here. The long internal branch to passerines may be attracting the long internal branch of parrots (see Hackett et al. 2008, fig. 3). At this point, support for such a grouping has not been found with mtDNA (e.g., Gibb et al. 2007; Brown et al. 2008). Cracraft (2001) by comparison included morphological, geographical, and early molecular data in support of his groupings. We therefore feel it is appropriate to use Cracraft (2001) as the basis for testing deep divergences within Neoavian birds rather than any one molecular data set.

As a step toward increasing the taxon sampling of coding sequences, we add nine new mt genomes: the dollarbird (*Eurystomus orientalis*) and New Zealand kingfisher (*Halcyon sancta vagans*) as representatives of the Coraciiformes, together with the white-tailed trogon (*Trogon viridis*) from Trogoniformes, are suggested to group with Piciformes; the great potoo (*Nyctibius grandis*) and Australian owl-nightjar (*Aegotheles cristatus cristatus*) as representatives of Caprimulgiformes; barn owl (*Tyto alba*), expected to pair with the New Zealand owl (morepork, *Ninox novaeseelandiae*) to form Strigiformes; the roadrunner (a ground cuckoo, *Geococcyx californica*) expected to pair with the New Zealand long-tailed cuckoo (*Eudynamis taitensis*) to form Cuculiformes; and the peach-faced lovebird (*Agapornis roseicollis*) from Psittaciformes expected to join with the budgerigar (*Melopsittacus undulatus*) and the ground parrot (kakapo, *Strigops habroptilus*). Thus, we have reduced the number of long branches in our data set by the addition of representatives from each of the six Neoavian lineages described in Cracraft (2001). For the third subgroup of the “Conglomerati”, only a rail (takehe) and the kagu are published. It is unclear whether these really are a natural group (Morgan-Richards et al.

2008), and this again leaves us with two long isolated branches. In accord with our previous practice, these species have temporarily been omitted until sequences from more closely related species are available for each of them.

We now change from the birds to the analysis. Perhaps, the most fundamental problem occurring while reconstructing deep-level phylogeny is substitution saturation (Curole and Kocher 1999; Phillips et al. 2006). Phylogenetic signal can be eroded by factors including superimposed substitutions and “nonhistorical” biases (such as from compositional nonstationarity)—which accumulate more rapidly at faster evolving sites. Attempts to limit these problems have been made in recent studies by identifying fast-evolving sites at which signal erosion is expected to be high (Morgan-Richards et al. 2008).

In previous work (Delsuc et al. 2003; Phillips and Penny 2003; Phillips et al. 2004), we found standard RY coding (Honeycutt and Adkins 1993), especially the third-codon positions, to be advantageous for the most variable partitions of nucleotide data. This recoding both increases the proportion of changes on internal branches of the tree (i.e., a “treeness” measure) and decreases the differences in nucleotide composition (relative compositional variability). This latter is important in reducing nucleotide composition effects because they have been long known to bias tree reconstruction (Lockhart et al. 1992). Because of the better fit of the data to the model (higher treeness and less variability in nucleotide composition), this has been our preferred method of analysis for vertebrate mt data.

Down-weighting the faster evolving sites or grouping faster evolving nucleotides (or amino acids) into a single category has been quite widely used (see Honeycutt and Adkins 1993; Philippe et al. 2000; Jeffroy et al. 2006)—although the theoretical aspects have not been well developed in phylogenetics. Rodriguez-Ezpeleta et al. (2007) report that omitting the fastest evolving sites, grouping amino acids into functional categories, and some mixture models, all enhanced the phylogenetic signal for deeper divergences. Susko and Roger (2007) similarly report improvements from down-weighting. However, some approaches may not be optimal if valuable sites are excluded simply because they are grouped under some prior definition (e.g., codon positions that have many fast-evolving sites). Conversely, some saturated sites may be retained because they are in a category that, on average, does not have site saturation. Thus, we can also group the justification for down-weighting into those using “a priori” categories (such as third position, stems vs. loops in RNA or amino acid groups) and “conditional” categories (down-weighting of each site independently).

We have used both in the past, the RY coding (a priori weighting category) (Phillips et al. 2004), and also a conditional down-weighting (Penny and Hendy 1986), based on the numbers of observed and expected incompatibilities. In general, all the methods mentioned here are examples of a standard statistical approach of “noise reduction/signal enhancement” (Proakis et al. 2002). Here, we implement a conditional noise reduction technique in which the information retained from the sequence is determined on a site-by-site basis. The Materials and Methods section has more

detail on this conditional recoding (down-weighting) of sites, an approach that we call site-stripping (Morgan-Richards et al. 2008).

Along with data partitioning/down-weighting and fossil calibrations (see Supplementary Material S1 online), our additional sequences mean that we can, in principle, calculate the probabilities that prior hypotheses are supported. In other words, we can calculate the proportion of trees that will have a split (or clade) that has been predicted. Unfortunately, there appears to be little use in phylogenetics for specifying a priori hypotheses and then testing the probability of finding them with new data. Rather, results are treated somewhat “post hoc,” looking at the trees after they are built and then trying to explain the results. In principle, a Bayesian approach allows alternative hypotheses to be given different weightings, but it appears that a “flat prior” is the norm; and this does not really differentiate between trees or hypotheses. Based on nuclear-coding sequences, Lin et al. (2002) took the four-way split within eutherian mammals and calculated the probability of finding the same split from mt data. The result was certainly very highly significant;  $P \approx 2.1 \times 10^{-7}$ . It is very important in phylogeny, though perhaps seldom carried out, to give quantitative estimates of the increase in information from a phylogeny. Penny et al. (1991) demonstrated that it is simple to calculate probabilities (see Appendix and table 1) of a single pair of taxa predicted to come together on the tree. That is, there is one chance in  $2n - 5$  of them coming together “by chance” on an unrooted tree and one in  $2n - 3$  for a rooted tree (given  $n$  taxa). It is even more improbable that a predefined grouping of three or more taxa will come together, and here, we develop measures in order to evaluate quantitatively the priors from Cracraft (2001).

Our approach here is to infer evolutionary trees from complete mt genomes as a start on resolving the deep Neoavian splits. We investigate the robustness of the priors set out by Cracraft (2001) by using a novel site-stripping method and exploring networks (see Supplementary Material S1 online for fossil calibrations). In addition, we assess quantitatively the usefulness of these priors by calculating clade probabilities. By finding resolution in the basal node of the Neoaves, hypotheses regarding the number of lineages present before the K–Pg boundary can be tested.

## Materials and Methods

### Taxon Sampling

The dollarbird (*E. orientalis*) and barn owl (*T. alba*) were supplied from the Australian Museum, Sydney, Australia, under sample numbers EBU 11118 and EBU 2564, respectively. The roadrunner (*G. californica*), great potoo (*N. grandis*), and the white-tailed trogon (*T. viridis*) were provided by the Louisiana State University Museum of Natural Science Collection of Genetic Resources under sample numbers LSUMZ B-8504, LSUMZ B-8954, and LSUMZ B-28495. The Australian owl-nightjar (*A. c. cristatus*) was provided by Fritz Geiser, University of New England, Armidale, Australia. The New Zealand long-tailed cuckoo (*E. taitensis*) and the New Zealand kingfisher (*H. s. vagans*) were obtained from Dick Gill, NZ Department of Conservation, Waikanae, New Zealand.

**Table 1**  
**Probabilities of a Predefined Clade Joining the Tree as a Single Group or as Two Subclades**

<i>k</i> Taxa in clade	Unrooted Trees Minimal	Rooted Trees Minimal
2	$1/(2n - 5)$	$1/(2n - 3)$
3	$3/(2n - 5)(2n - 7)$	$3/(2n - 3)(2n - 5)$
4	$15/(2n - 5)(2n - 7)(2n - 9)$	$15/(2n - 3)(2n - 5)(2n - 7)$
5	$105/(2n - 5)(2n - 7)(2n - 9)(2n - 11)$	$105/(2n - 3)(2n - 5)(2n - 7)(2n - 9)$

The peach-faced lovebird (*A. roseicollis*) was obtained locally from commercial breeders.

# Molecular Methods

Extractions of genomic DNA from each of the newly sampled birds were performed at the Allan Wilson Centre from 25 to 50 mg of liver tissue using the High Pure PCR Template Preparation Kit (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. To minimize the chance of obtaining nuclear copies of mt genes (NUMTs), 2–4 overlapping long-range polymerase chain reaction (PCR) fragments (3.5–12 kb in length) were first amplified using the Expand Long template PCR System (Roche Applied Science). The products were excised from 1% agarose gels and purified using a QIAquick Gel extraction kit (Qiagen GmbH, Hilden, Germany) as per the manufacturer's instructions. These long-range products were subsequently used as template DNA for following short-range PCRs (overlapping fragments 0.5–3 kb in length). Short-range primer combinations were found using our laboratory database as described in Slack et al. (2006), and any new primers required were designed using Oligo 4.03 (National Biosciences, Inc., Plymouth, MN). Sequencing was performed using BigDye Terminator Cycle Sequencing reagents according to the manufacturer's instructions (Applied Biosystems, Foster City, CA) and then sequenced on an ABI 3730 automated sequencer (Applied Biosystems). Sequences were aligned using Sequencher 4.7 (Gene Codes Corp., Ann Arbor, MI) and then manually edited and checked for complete concurrence between overlapping sequences.

Where necessary (e.g., with length heteroplasmy in control regions [CRs] from microsatellite repeats), PCR products were cloned using the TOPO TA cloning kit for sequencing (Invitrogen, Carlsbad, CA). For each region, at least three clones were sequenced to safeguard against PCR errors. In all cases, overlaps between sequences were sufficient to ensure synonymy and sequence identity was confirmed through Blast searches (<http://www.ncbi.nlm.nih.gov/blast/>), confirmation of amino acid translation in coding regions, and alignment with other species.

In addition to the nine new bird mt genomes reported in this paper, 36 other complete avian mt genomes from NCBI GenBank were included in the analyses: 31 Neoaves and 5 Galloanserae. Paleognath taxa were not included in this data set; although their overall placement is now well established (Gibb et al. 2007; Slack et al. 2007), there are still important but unresolved issues around the placement of tinamous (Hackett et al. 2008; Harshmann et al. 2008; Phillips MJ, Gibb GC, Crimp EA, Penny D, in preparation). Instead, we rooted our Neo-

aves trees with the Galloanserae sequences (Gibb et al. 2007; Morgan-Richards et al. 2008). The full data set is available from the authors on request.

The Galloanserae taxa are Japanese quail (*Coturnix japonica*, AP003195), magpie goose (*Anseranas semipalmata*, AY309455), redhead duck (*Aythya americana*, AF090337), greater white-fronted goose (*Anser albifrons*, AF363031), and Australian brush turkey (*Alectura lathamii*, AY346091). The 31 Neoaves taxa (modern birds) are rifleman (New Zealand wren, *Acanthisitta chloris*, AY325307), gray-headed broadbill (*Smithornis sharpei*, AF090340), fuscous flycatcher (*Cnemotriccus fuscatus*, AY596278), superb lyre bird (*Menura novaehollandiae*, AY542313), rook (*Corvus frugilegus*, Y18522), ivory billed toucan (*Pteroglossus azara*, DQ780882), pileated woodpecker (*Dryocopus pileatus*, DQ780879), morepork (a New Zealand owl, *N. novaeseelandiae*, AY309457), kakapo (flightless parrot *S. habroptilus*, AY309456), budgerigar (*M. undulatus*, EF450826), ruby-throated hummingbird (*Archilochus colubris*, EF532935), common swift (*Apus apus*, AM237310), peregrine falcon (*Falco peregrinus*, AF090338), forest falcon (*Micrastur gilvicolis*, DQ780881), Eurasian buzzard (*Buteo buteo*, AF380305), osprey (*Pandion haliaetus*, DQ780884), Blyth's hawk eagle (*Spizaetus alboniger*, AP008239), blackish oystercatcher (*Haematopus ater*, AY074886), ruddy turnstone (*Arenaria interpres*, AY074885), southern black-backed gull (*Larus dominicanus*, AY293619), red-throated loon (*Gavia stellata*, AY293618), little blue penguin (*Eudyptula minor*, AF362763), rockhopper penguin (*Eudyptes chrysocome*, AP009189), black-browed albatross (*Diomedea melanophrys*, AY158677), Kerguelen petrel (*Pterodroma brevirostris*, AY158678), frigatebird (*Fregate* sp., AP009192), Australian pelican (*Pelecanus conspicillatus*, DQ780883), Australasian little grebe (*Tachybaptus novaehollandiae*, EF532936), greater flamingo (*Phoenicopterus ruber roseus*, EF532932), great crested grebe (*Podiceps cristatus*, AP009194), and the Oriental white stork (*Ciconia boyciana*, AB026193).

# Phylogenetic Analysis

Sequences were aligned in Se-Al v2.0a11 at the amino acid level for protein-coding genes and based on secondary structure for RNA genes. The data set has 12 protein-coding genes, 2 ribosomal RNAs (rRNA), and 22 transfer RNAs (tRNA). Gaps, ambiguous sites adjacent to gaps, NADH6 (light-strand encoded), and stop codons (often incomplete in the DNA sequence) were excluded from the alignment. The 12 protein-coding genes were separated into first-, second-, and third-codon positions, whereas rRNA and tRNA genes were partitioned into stems (S) and loops (L).

Protein-coding genes were checked for NUMTs by translating into amino acids.

Previous studies from birds (Slack et al. 2003), mammals (Lin et al. 2002), and simulations (Holland et al. 2003) have all shown that the addition of outgroups can disrupt the ingroup tree. However, in such cases (from theory and with simulated data), the ingroup tree (i.e., with the outgroup omitted) is more likely to be correct (Holland et al. 2003). We therefore ran separate analyses either including or excluding the outgroup (five birds from the Galloanserae). A combined total of 13,412 nucleotides (excluding gaps) were used for the basis of further analyses (see below for number of characters per data set). As mentioned earlier, we partitioned the data: codons 1 and 2, codon 3, RNA stems, and RNA loops (Slack et al. 2007) for site-stripping. ML analyses were carried out using standard programs including PAUP\* 4.0b10 (Swofford 2001) and GARLI v0.95 (Zwickl 2006). Bayesian analysis was carried out in MrBayes (Huelsenbeck and Ronquist 2001), and consensus networks were implemented in SplitsTree version 4 (Holland et al. 2004; Huson and Bryant 2006) and BEAST (Drummond et al. 2006; Drummond and Rambaut 2007). Optimal parameters for the ML models were determined using Modeltest 3.7 (Posada and Crandall 1998) and the AIC values used. The hierarchical and AIC tests were in agreement for Modeltest. Initial results from ML analyses were consistent with Bayesian analysis and were not used for site-stripping. The best model for the full data set was GTR + I + G. Bayesian analyses used default parameters and were run for 10 million generations or until convergence was obtained. In addition to Bayesian posterior probabilities (BPPs), we ran analyses in PhyML (Guindon and Gascuel 2003) and RAXML (Stamatakis 2006) to carry out 100 bootstrap replicates on the data sets both with and without the outgroup (see Supplementary Material S2 online for results).

#### Noise Reduction by Down-weighting (Site-Stripping)

Site-stripping compares sites based on the actual number of mutations required on the tree ("tree steps") versus the maximum possible number of mutations for that site (max). The calculation is the limit ( $L$ ) = tree steps  $\times$  tree steps/max steps. If the threshold "strictness" ( $s$ ) is, for example, 4, then sites for which  $L \leq 4$  remain unchanged and sites for which  $L > 4$  are RY coded. If after RY coding  $L$  is still  $> 4$ , the site is excluded. The higher the threshold (the larger the value of  $s$ ), the more sites that are included and fewer are RY coded. Conversely, the lower the threshold (lower values of  $s$ ), the more sites RY coded or excluded. Therefore, the weighting of each site is a function of that site and is not predetermined by being a member of a class. This allows, for example, some hypervariable sites from first and second positions to be RY coded or omitted. A range of  $s$  values were used and resulted in the following:

$s = \infty$ , 0 sites RY coded, 0 sites excluded (all 13,412 sites included);

$s = 6.0$ , 20 sites RY coded, 0 sites excluded;

$s = 4.5$ , 91 sites RY coded, 1 site excluded;

$s = 3.5$ , 254 sites RY coded, 8 sites excluded;

$s = 2.28$ , 578 sites RY coded, 63 sites excluded; and

$s = 2.0$ , 891 sites RY coded, 159 sites excluded.

Bayesian inference analyses were carried out on each of the data matrices, including the fully weighted data. Note that RY coding increases the ML scores as it amalgamates some nucleotide categories; thus, the data are now different, and it is not valid to compare directly ML scores from RY and nucleotide coding (Steel MA, personal communication). Similarly, consistency index (CI) values are not directly comparable between nucleotide and RY coded data sets. That is, for an unrooted four-taxon tree, for which A, C, T, and G states are random (i.e., no signal remains), the expected average CI for nucleotide data is 0.949, whereas for the RY coded data, it is 0.778.

In order to identify fast-evolving sites and so facilitate noise reduction by site-stripping, nine additional close relatives were added to the alignment but were removed before phylogenetic analysis as they do not break up long branches or add phylogenetic signal and would only increase analysis times. The nine taxa included were chicken (*Gallus gallus*, AP003317); green junglefowl (*Gallus varius*, AP003324); gray junglefowl (*Gallus sonneratii*, AP006741); white stork (*Ciconia ciconia*, AB026818); Canadian goose (*Branta canadensis*, DQ019124); tundra swan (*Cygnus columbianus*, DQ083161); mountain hawk eagle (*Nisaetus nipalensis*, AP008238); Pacific loon (*Gavia pacifica*, AP009190); and American kestrel (*Falco sparverius*, DQ780880) (see Supplementary Material S3 online for the consensus tree of the above nine taxa and their close relatives).

#### Probabilities of Observing Predefined Clades

It is straightforward to calculate the probability of observing a prespecified clade in a tree from a new data set, that is, a tree using additional data not used to predict the clade (see Appendix A for details). In general, for  $n$  taxa, there are  $B(n) = (2n - 5)!!$  unrooted binary trees, where the double factorial notation (!!) is the product of every second number, that is,  $1 \times 3 \times 5 \times \dots \times 2n - 5$  (table 1). Thus, for example, there are  $\approx 3 \times 10^{20}$  possible unrooted binary trees for 20 taxa. In addition, even if a predicted clade for three or more taxa ends up as two subclades on the tree, we can calculate the probability of observing this; and it can indicate that there is still high information content in the data. The calculation can be extended further to three or more subclades, but here, we concentrate mainly on our prior clades being found. We need to define the composition of the prior clades carefully; they alter with the question being considered. For example, we may be interested whether the three parrots do form a natural group, in which case we would calculate the probabilities of three parrots coming together with  $n = 40$  taxa in the data set—this is close to a flat prior. Alternatively, we may accept the grouping within orders and consider just  $n = 21$  deep Neoavian groups (see later, fig. 2); this is testing the groupings ("priors") of Cracraft (2001). Other tests are also possible.

#### Results

The first of our three approaches to improving the Neoavian tree was the inclusion of more sequences. The

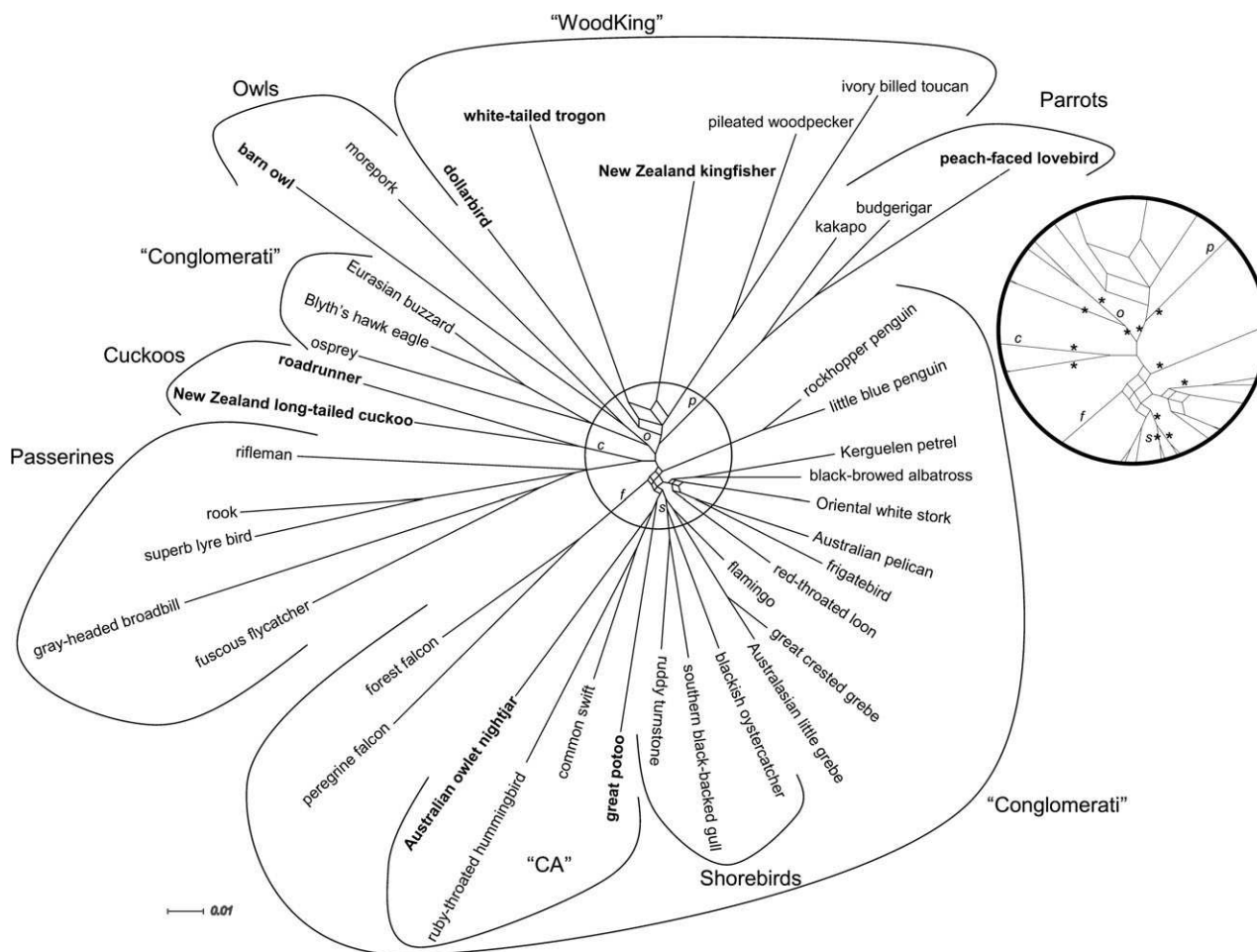


FIG. 1.—Unrooted Bayesian consensus network of Neoaves (modern birds), based on whole mtDNA genomes with “minimal” down-weighting (threshold  $s = 6.0$ ; 20 sites RY coded, 0 sites excluded). Only splits occurring in  $>25\%$  of trees are included in the network. Inset shows the central portion (indicated by the circle) expanded. Selected branches are labeled to ease comparison: c, cuckoos; f, falcons; o, owls; p, parrots; and s, shorebirds. New taxa included are highlighted in bold. Splits indicated by asterisk have 99+ Bayesian posterior support (BPP).

nine new mt genome sequences are deposited in GenBank under the following accession numbers: dollarbird (EU344978, 17,774 bp); barn owl (EU410491,  $>16,148$  bp, incomplete ND6 and CR); roadrunner (EU410488, 17,091 bp); great potoo (EU344977,  $>14,396$  bp, incomplete ND6 and CR); white-tailed trogon (EU410490, 17,751 bp); Australian owl-nightjar (EU344979, 18,607 bp); peach-faced lovebird (EU410486, 16,732 bp); New Zealand long-tailed cuckoo (EU410487, 17,559 bp); and the New Zealand kingfisher (EU410489, 17,549 bp). Following the gene-order nomenclature from Gibb et al. (2007), the cuckoo and roadrunner were found to have the remnant CR(2) gene order first described in the falcon (Mindell et al. 1998). All other birds have the standard avian gene order as in the chicken (Desjardins and Morais 1990) with the possible exception of the great potoo and barn owl; because their sequences are currently incomplete in the CR and adjacent genes, their gene order is unknown.

Site-stripping (noise reduction) was our second approach to improving the tree. Bayesian analyses were carried out on a range of down-weighting values, initially excluding the outgroup, then including it. Figure 1 shows

the result for the unrooted Neoavian data set with a high threshold (minimum down-weighting, strictness  $s = 6.0$ ), whereas figure 2 shows results from the maximum down-weighting used (a low threshold, strictness  $s = 2.0$ ). Both figures are networks showing splits occurring in at least 25% of Bayesian phylogenies (Holland et al. 2004).

Our third approach was a quantitative test of predefined groupings. There are several clades that were predicted and were returned both with different down-weightings and with or without the outgroup. Perhaps, the most straightforward example is that of the three parrots; this is just a trial calculation because we really had no doubt that the parrots would come together, as predicted from previous DNA sequence analyses (de Kloet RS and de Kloet SR 2005). The New Zealand ground parrot (kakapo) was the first to diverge. This is just a trial calculation with flat priors, and the probability of three taxa coming together on a tree is  $P = 0.0005$  ( $P = 3/(2n - 5)(2n - 7)$ , with  $n = 40$  (see Appendix A). There is considerable rate variation within parrots, with the kakapo being slower than the others, whereas the peach-faced lovebird is the

fastest (which is evident by the long edge in both figs. 1 and 2).

Again as expected from our informal priors, the New Zealand long-tailed cuckoo and the roadrunner (a ground cuckoo) always paired and do in prior morphological and molecular studies; thus, we considered them to form one independent “taxonomic group.” Of the 40 taxa analyzed, we consider only 21 groups to be independent (see later and Appendix A for details of the groupings); consequently, the probability of two taxa forming a clade in the tree is  $P = 0.027$  ( $P = 1/(2n - 5)$ , with  $n = 21$  see Appendix A). Perhaps unexpectedly, cuckoos then group as sister to the five passerines in our data set, a result also observed by Mayr et al. (2003) using combined molecular and morphological data; however, bootstrap support was relatively low (see their figs. 5 and 7). As in previous analyses, the passerines always group together, with the New Zealand wrens (rifleman in this case), basal to the oscines and suboscines. The cuckoo–passerine pairing was found with all down-weightings, both with the ingroup alone and with the outgroup included (fig. 3). This grouping of the cuckoo–passerine clades is an interesting hypothesis and requires testing with both nuclear-coding sequences and rare genomic changes (Boore 2006).

The first real test of Cracraft’s (2001) priors, that is, testing groups above the order level, stems from the clade we refer to as “WoodKing”. In this case, all five taxa were always found as a clade, irrespective of the down-weighting, and with or without the outgroup (i.e., with both unrooted and rooted trees). As expected, the Piciformes (pileated woodpecker and ivory billed toucan) were always paired. Thus, if it is assumed that the woodpecker and toucan are sufficiently close, then really there are only four independent taxa, and  $P = 7.12 \times 10^{-6}$  is a highly significant result (see Appendix A for details). Note that the calculation of the probabilities allows all possible ways of observing the four taxa on the tree, including any paraphyletic groups within it. As shown in figures 1 and 2, there is conflicting signal linking the kingfisher with either the white-tailed trogon or with the dollarbird. Being able to show both signals is a major advantage of networks (Holland et al. 2004) because it helps to prevent premature conclusions. With increased down-weighting of the faster sites (a stricter threshold), we again observed variation in the position of the dollarbird, which tended to be deeper in the clade, though still within the “WoodKing” group. Additional taxon sampling should resolve the splits fully, but our main conclusion is that the predicted grouping of Piciformes, Trogoniformes, and Coraciiformes (Cracraft 2001) is found (though not necessarily reciprocally monophyletic). Hackett et al. (2008) sampled more widely and found strong support for Coraciiformes + Piciformes (see their fig. 2, clade C), although Trogoniformes fell outside this and had less support (see our Supplementary Material S2 online).

However, the next result was not in our informal priors. We found that the parrots and the WoodKing group, irrespective of down-weighting extent, are always adjacent clades on the unrooted tree (supported by high BPP values but not by bootstrapping, see Supplementary Material S2 online). This result will need further investigation as the parrot lineage has considerable rate variation and there is

a long internal branch from the three parrots to the rest of the tree. Because the grouping was not part of our priors, we cannot calculate the increase in support, but if additional data types support this relationship, then the probabilities could be calculated.

Next, we consider the “SCA” group (Strigiformes, Caprimulgiformes, and Apodiformes) predicted by Cracraft (2001). Four of the six taxa available form a monophyletic clade, to the exclusion of the two owls. The great potoo, Australian owl-nightjar, common swift, and ruby-throated hummingbird formed a group of four (“CA”). The swift and hummingbird pairing was highly supported by BPP and bootstrapping (see Supplementary Material S2 online) as predicted from previous studies (Johansson et al. 2001; van Tuinen and Hedges 2001; Cracraft et al. 2004; Hackett et al. 2008; Morgan-Richards et al. 2008), the Australian owl-nightjar came deeper, and finally the great potoo, which was always basal. It should be noted that the potoo did move slightly with bootstrapping, and we found more support for the Apodiformes and the owl-nightjar to the exclusion of the potoo (see Supplementary Material S2 online) (see also Mayr 2002a; Barrowclough et al. 2006). Hackett et al. (2008) also observed high support (98%) for Apodiformes + *Aegotheles*, as have previous nuclear and morphological studies (see also Mayr 2002b; Barrowclough et al. 2006). The position of this group of four was variable in the tree. With lesser down-weighting (a higher threshold,  $s = 6.0$ ), the group was found within the informal “Conglomerati” group (fig. 1), but with the maximum down-weighting, it was outside this group (apart from the shorebirds). Similarly, the barn owl and the morepork always paired, joining together quite deep in the tree (i.e., although both are “owls,” they represent old divergences; indeed some short preliminary runs did not even join them together, see Supplementary Material S2 online for support values). We need to be cautious here as both owls have some of the highest rates of sequence evolution among the Neoaves. Although the owls did not group with the other members of the “SCA” clade, they were always found to group with the buzzard, hawk eagle (Accipitridae), and osprey. Only with the highest down-weighting used (figs. 2 and 3) did the falcons unite with the owls/Accipitridae/osprey; with the falcons basal, however, we found no bootstrap support and low BPP support for this grouping (see Supplementary Material S2 online). Grouping of the Accipitridae, osprey, and owls is certainly interesting. Relationships within the birds of prey are controversial as speculation over convergence and raptorial specialization has been raised (e.g., Livezey and Zusi 2007). This relationship was not found by Hackett et al. (2008) and needs to be tested with nuclear-coding data. We return to the raptors and owl question later.

Even though the proposed “SCA” clade came out as two groups in the tree, this in itself still has high information content. The simplest calculation is to assume the alternative prediction of the two owls being independent of the other four, and in this case, the probability of observing the “two plus four” grouping is simply the product of the probability of observing a pair ( $P = 1/(2n - 5) = 0.013$  for  $n = 40$ ) and the probability of observing a group of four ( $P = 2.5 \times 10^{-6}$ , see Appendix). This gives the combined

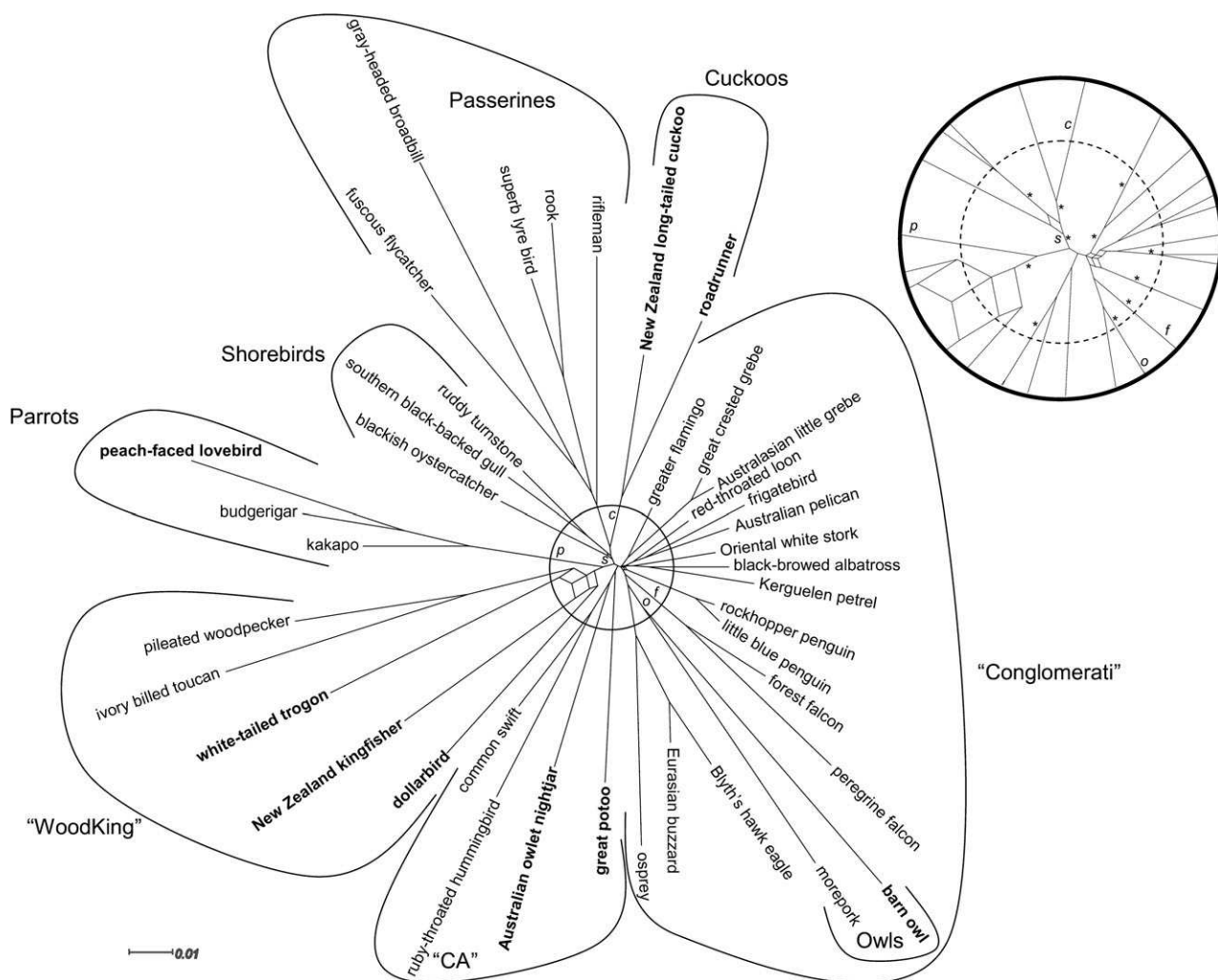


FIG. 2.—Unrooted Bayesian consensus network of Neoaves with the “maximum” down-weighting used (threshold  $s = 2.0$ ; 891 sites RY coded, 159 sites excluded). Otherwise, the conventions are the same as in figure 1.

probability of  $3.4 \times 10^{-8}$  (about 1 chance in 300 million of observing the pattern). Strictly speaking, we need not make the assumption that the two owls (in particular) separate from the other four taxa; there are a total of 21 combinations ( ${}^6C_2$ ) of pairs from six taxa; only one has the two owls paired. In general, we would multiply the probability by 21 (see Appendix A), still giving around 1 chance in 10 million of observing these two groups on the tree. Similarly, it could have been just a single taxon (six choices) that separated from the other five or two groups of three taxa (10, or  ${}^6C_3$  combinations) but halved because each triplet of taxa is found twice. Appendix A and figure 4 show the general calculation, but in this case (because the owls join with the raptors, as an alternative prediction see Mayr 2005), it is reasonable to use the probability of finding just the two plus four grouping.

Our conclusion at this point is that excellent progress is being made in understanding the deeper levels of phylogeny of the Neoaves. If the predictions from Cracraft (2001) are, in general, being well supported with new data, then this implies that the basal polytomy is resolvable. However, our next step is to check that there are no major changes when the outgroup is added; this has been a major problem

when a smaller number (24 ingroup) of taxa were sequenced (see Harrison et al. 2004).

### Rooted Tree

Figure 3 shows our tree rooted with five Galloanserae taxa for the maximum down-weighting value (threshold  $s = 2$ ). The two main points from this figure are that

1. there is only one local change to the unrooted tree when the root is added, and
2. the root comes between parrots and all other Neoaves.

The first point refers to the difference in the position of the flamingo/grebe clade between the unrooted and rooted trees. In the unrooted tree, they are basal to the loon/albatross/Pelecaniformes/stork grouping. In the rooted tree, we find them basal to the same group as before plus the Accipitridae/osprey/owl/falcon clade. In other words, the tree is “locally stable” in the terminology of Cooper and Penny (1997). Finding just a single edge (branch) different between two trees is very highly significant, about  $5.7 \times 10^{-54}$  for 40 taxa in the ingroup (Penny et al. 1982).

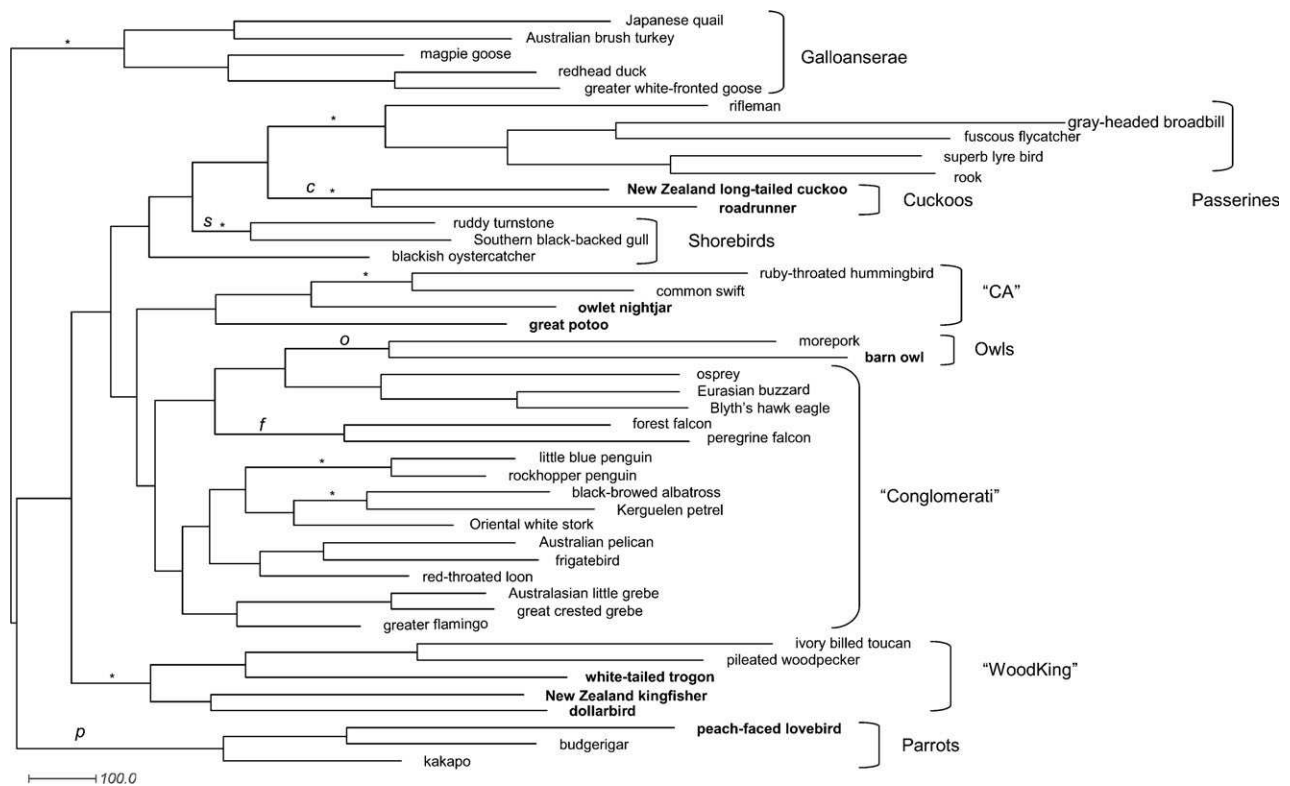


FIG. 3.—Rooted Bayesian phylogram of Neoaves with maximum down-weighting and including the five Galloanserae as the outgroup. Conventions are the same as in figure 1.

On the surface, this latter finding (the root between parrots and other Neoaves) could be suspicious because the branch at the base of the parrots is the longest internal branch on the tree! However, with lesser down-weighting ( $s = 6.0$ ), the root joins one step away on the short branch at the base of the parrot/“WoodKing” group. The movement of the root from a shorter branch is not expected unless parrots are really the first subdivision of Neoaves. Without any prior information as to the root of the Neoaves, it is not possible to give any quantitative statement of confidence in this rooting.

Other aspects of the rooted tree are also interesting and noteworthy. We again find the same owl/raptor clade appearing only with the strongest down-weighting—that is, only the Accipitridae/osprey (but not falcons) unite with the owls with lesser down-weighting. The tree produced from BEAST resulted in the falcons joining the “CA” group (Supplementary Material S1 online). This latter observation, if real, is interesting as both falcons and the “CA” taxa are in-flight foragers typically specialized for bill-capture of prey; in comparison, the owls are specialized for in-flight talon capture (like hawks). However, there was only conflicting support for this relationship shown in the network figure 1, and therefore, we are unable to comment further without additional data. It is worth noting that some of the deeper groupings within this large raptor/shore bird/water bird (“Conglomerati”) still vary somewhat depending on taxon sampling (Morgan-Richards et al. 2008), and it is not clear yet whether further taxon sampling will resolve these issues. It may well be that the three subgroups in the

“Conglomerati” (group (vi) of Cracraft 2001, see Introduction) should be considered independently.

## Discussion

Resolving the evolutionary relationships within the modern birds (Neoaves) has been both problematic and controversial, with some suggesting that it will never be fully resolved (Poe and Chubb 2004). Here, we have shown that with more and longer DNA-coding sequences, along with improved noise reduction techniques, relationships within Neoaves are expected to be resolvable. This should occur relatively quickly with the addition of data from nuclear coding and rare genomic changes as they become available. Our approach to resolving this issue has been 3-fold: presenting additional data, improving noise reduction/signal enhancement techniques, and getting beyond flat priors, where it is assumed (sometimes correctly!) that there is no useful prior knowledge.

We consider it important that prior hypotheses can be evaluated quantitatively, and thus, the formulae developed in Appendix A will be useful for a wide range of studies. However, there is still more work required in developing these analyses. For example, the calculation for prespecified groups is for the optimal placement of that clade on the tree (even if the bootstrap or Bayesian posterior priors are less than 100% support). If these support values are indeed higher, then this gives even more confidence in the clades, so in that respect our probabilities are conservative. Thus, more thought is required on how to combine the

calculations developed here with the strength of support for branches in the tree from new data. Similarly, the calculation allows for any subtree within the clades (or subclades). However, if we prespecified that a particular grouping and subtree is expected, then the probability of finding this arrangement on the tree is even lower. For example, the calculation allows 15 ways (5!!) that a group of four could join the larger tree. This is because there are three unrooted trees for four taxa, each with five edges for joining to the rest of the tree—and the prediction does not specify which of the  $3 \times 5$  (15) trees would be observed. In contrast, if we predict precisely how the group will join (forming a clade), then the number of possible trees is reduced 15-fold. Overall, it is important that we make better use of well-considered prior hypotheses when studying trees based on new data. Even though we are a long way from having the “one tree” (in this case) for Neoaves, we can be confident that the issue is resolving and that the data sets have, in a formal sense, high information content.

A related question is estimating how many trees within Neoaves are still likely—a “confidence set” of trees. At this higher taxonomic level, it is not yet clear which groupings are stable and which may be subject to change. For our Neoaves data set, there appears to be around 21 major groupings (shown as crossing into the inner dashed line in the insert of fig. 2 and in Appendix A). These groupings include Passeriformes, cuckoos, parrots, two shore bird lineages, three raptor clades (falcons, buzzard/osprey, and owls), rollers, kingfishers, woodpecker/toucan/trogon, potoos, owl-nightjars, hummingbirds/swifts, flamingoes, grebes, Pelecaniformes, tubenoses, storks, penguins, and loon.

In principle, there are  $R(21) \approx 3 \times 10^{23}$  possible rooted binary trees of which only a vanishingly small proportion are realistic. With eutherian mammals, there were initially 19 orders identified and therefore  $R(19) \approx 2 \times 10^{20}$  possible rooted binary trees. But it quickly became apparent that no more than about  $10^2$  trees were likely (Lin et al. 2002)—an improvement of 18 orders of magnitude. The next step for birds is an equivalent analysis for Neoaves and thus getting beyond the debilitating view of the flat priors—that all trees were equally likely and that there is no information in previous studies.

The results from down-weighting the faster evolving sites are interesting, and these techniques need to be developed and tested further. With increased down-weighting, we find closer agreement between earlier predictions and the actual tree found. From first principles, we expect that reducing the influence of the saturated sites will help, and in general, it appears that the predefined groups are found more strongly. It is for this type of reason that we would like to see further development and evaluation of the noise reduction techniques including their application to nuclear-coding data. Although it is outside the range of this study, a simulation study is now an important next step.

Turning now from the more general issues to the Neoaves in particular, our current study assesses the stability and probability of the six groups proposed by Cracraft (2001) using a novel analysis method to down-weight sites of whole mt genome sequences. These Cracraftian priors were found to be robust with four of the six groups within Neoaves being recovered, the other two having relatively small changes; the

owls moving to the raptors and possibility that the raptor/water carnivores (“Conglomerati”) may be diphyletic. In this latter respect, it appears preferable, at least in the short term, to treat the three subgroups of Cracraft’s group (vi) independently. Our resulting phylogenies appear relatively stable, differing little in overall topology both with and without the addition of the outgroup (Galloanserae).

The placement of the root of Neoaves needs additional support. Our analyses put the root in one of two possible locations: either with parrots (which have a higher mutation rate in mtDNA) as the most basal lineage or, with lesser down-weighting, the parrots plus the “WoodKing” grouping as the basal clade. Because the separation of parrots from the rest of the Neoaves occurred with the highest down-weighting, we cannot easily dismiss this possible rooting. Morphologically, parrots are distinct (for review, see Dyke and Cooper 2000; Waterhouse 2006) and a fragment of a mandible from the Maastrichtian, latest Cretaceous (65–70 Ma Lance Formation, North America) has been described (Stidham 1998). However, the identification of this fossil is contentious (Dyke and Mayr 1999; Mayr 2002a), though previous molecular work suggests a Cretaceous diversification for each of the African, Australian, and South American parrots (Miyaki et al. 1998). Dating carried out on our current data set suggests that the most basal parrot in our analysis, the kakapo, split from the other parrots sometime after the K–Pg boundary. The lineage as a whole however predates the K–Pg boundary with a mean date of  $\sim 85$  Ma (see Supplementary Material S1 online). Hackett et al. (2008) suggested that the root of Neoaves be placed with the sister grouping of the Podicipediformes/Phoenicopteriformes/Phaethontidae/Pteroclididae/Mesitornithidae/Columbiformes to Gruiformes/Caprimulgiformes/Apodiformes (see their fig. 2, clades K, L, M, and N). However, they do not give support values for this lineage, and they state that their rooted tree only occurs when the  $\beta$ -fibrinogen intron 7 data are included. We have already demonstrated (Morgan-Richards et al. 2008) that by our more rigorous standards (deleting columns around gaps back to a constant column) that the intron sequences of this locus are not informative for deep divergences. Clearly, the root of Neoaves is still under debate; however, we now have a number of possibilities to be tested by future analyses.

## Conclusion

The basal split within Neoaves and its timing are resolvable issues. If modern birds radiated over a short period (say 2–5 Ma) after the K–Pg extinction, then it will be very difficult to resolve the polytomy at the base of modern birds. However, using whole mtDNA coding sequences gives us a solid point from which to build. By the addition of more taxa, nuclear-coding sequences, and rare genomic changes, we expect resolution at the ordinal level to be achievable. In addition, the further development of noise reduction techniques for coding sequences (both organellar and nuclear) will enable more robust trees to be produced. We estimate that at least 12 Neoavian lineages had evolved prior to the K–Pg boundary, similarly, van Tuinen et al. (2006) and Brown et al. (2008) support pre-K–Pg origins

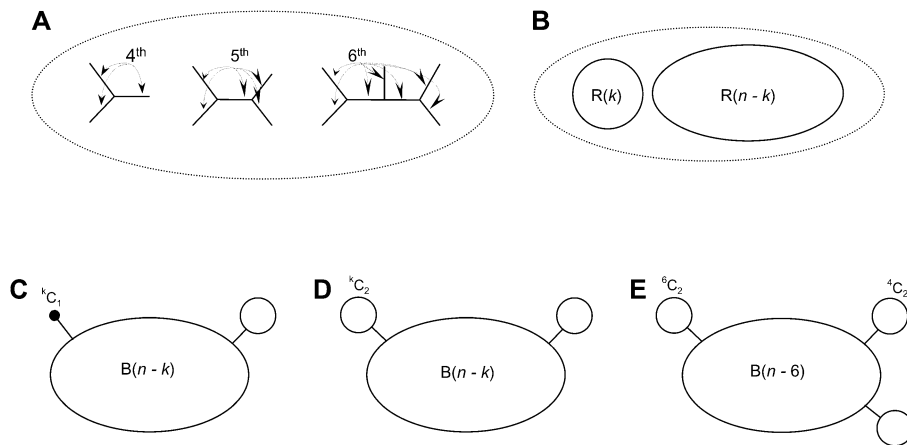


FIG. 4.—The basis for calculations for the probability of finding predefined clades (or subclades) on a tree. (A) A procedure for counting the number of trees. There is only one unrooted tree for three taxa, and there are three edges (branches) to add the fourth taxon—giving three trees for four taxa. Each of these three trees has five edges where the fifth taxon can be added giving  $1 \times 3 \times 5 = 15$  trees. Similarly, each of these 15 trees has 7 edges for the sixth taxon to be added, leading to the formula  $B(n) = (2n - 5)!!$ , for the number of unrooted binary trees. (B) Calculating the probability of a prespecified clade of  $k$  taxa on a tree. There are  $R(k)$  rooted trees for the clade of  $k$  taxa and  $B(n - k)$  for the remaining  $n - k$  taxa, leading to the calculation for the probability of observing a prespecified clade of  $k$  taxa forming a clade in each tree with  $n$  taxa. (C) The proposed grouping of  $k$  taxa with two subgroups; a single taxon in one and  $k - 1$  in the other. (D) A similar case with two taxa on one group and  $k - 2$  in the other. There are  ${}^kC_2$  ( $k$  choose 2) ways of selecting the two taxa. (E) An example where the group of  $k = 6$  ends up as  $m = 3$  subgroups of 2 taxa each. There are  ${}^6C_2$  for selecting the first pair of taxa and  ${}^4C_2$  for the second pair and  $3!$  ways for ordering the three pairs on a given tree.

for multiple modern lineages. In addition, Clarke et al. (2004) estimate a minimum of five Anseriformes lineages (duck, chicken, and ratite bird relatives) before this time supporting the presence of a diverse array of modern bird lineages prior to the extinction event. Lastly, with regards to the search for the one tree, we feel our data have made significant progress with support for four of the six Cracraftian groups. Given the very low probability of observing groupings by chance, the data are highly informative and should stimulate future work incorporating data from all facets of avian evolution.

### Supplementary Material

Supplementary Material S1 (fossil calibrations and molecular dating with BEAST), Supplementary Material S2 (tabulated results from Bayesian analysis [BPP] and bootstrap analyses [bs] using RAXML and PhyML), and Supplementary Material S3 (consensus tree including all outgroup taxa used for alignment before for site-stripping) are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

### Acknowledgments

We thank the Australian Museum, Sydney; the Louisiana State University Museum of Natural Science Collection of Genetic Resources; Fritz Geiser, University of New England, Armidale, Australia; and Dick Gill, Department of Conservation, Waikanae, for supplying tissue samples. Thanks to Trish McLenachan for help with alignments and sequences and to Olga Kardailsky and Steve Trewick for additional mitochondrial genomes. Thanks to Simon Hills and Patrick Biggs for computer time and assistance with BEAST analyses. Work carried out at the Allan Wilson Centre for Molecular Ecology and Evolution was

supported by the Marsden Fund (Royal Society of New Zealand) to D.P. The manuscript was improved by constructive comments from Associate Editor M. Gouy and three anonymous referees.

### Appendix A

#### The Probability of Observing a PreSpecified Clade

We calculate the probability of observing a prespecified clade of  $k$  taxa in a binary tree on  $n$  taxa, as the proportion of all binary trees containing that clade. This can be extended to the probabilities of the clade being found as  $m = 2, 3$ , or even more subclades. We see below (that for all but small values of  $n$  and  $k$ ) that these probabilities are very low so that finding a prespecified clade on a tree formed from new data is highly significant.

It is well known that for  $n$  taxa there are

$$B(n) = (2n - 5)!! = 1 \times 3 \times 5 \times \dots \times (2n - 5) \quad (1)$$

unrooted binary trees where each tip (leaf) of the tree is labeled by a unique taxon (see Penny et al. 1991). Similarly, the number of rooted binary trees is

$$R(n) = (2n - 3)!! = B(n + 1). \quad (2)$$

A simplified approach to deriving the formulae is indicated in figure 4A, and the calculations are straightforward in an Excel spreadsheet.

#### Probability of a Specific Subset Forming a Single Clade (i.e., $m = 1$ )

The probability ( $P$ ) of observing a predefined clade of  $k$  taxa in a binary tree of  $n$  taxa is

$$P(n, k) = R(k) \times B(n - k + 1)/B(n), \quad (3)$$

where the numerator is the number of rooted subtrees for the clade ( $R(k)$ ), multiplied by the number of trees on the remaining taxa ( $B(n - k + 1)$ , including a leaf for the clade). Dividing by the number of unrooted trees ( $B(n)$ ) gives the proportion of trees having that clade of  $k$  taxa.

For rooted trees, the probability ( $P_R$ ) of observing a predefined clade of  $k$  taxa in a rooted binary tree of  $n$  taxa is similarly

$$P_R(n, k) = R(k) \times R(n - k + 1)/R(n) = P(n + 1, k). \quad (4)$$

For this question, we consider all binary trees as equally likely; the trees are derived from a Markov model where there is no prior information about the distribution of tree shapes (Steel and Penny 1993). For two taxa in the predefined clade, the equation simplifies (see table 1) to

$$P(n, 2) = 1/(2n - 5) \text{ for unrooted trees and}$$

$$P_R(n, k) = 1/(2n - 3) \text{ for rooted trees.}$$

And for three taxa, it simplifies to

$$P(n, 3) = 3/(2n - 5)(2n - 7) \text{ for rooted trees and}$$

$$P_R(n, 3) = 3/(2n - 3)(2n - 5) \text{ for rooted trees.}$$

In our analyses, we consider the 40 taxa to account for 20 independent taxonomic groupings: passerines (five taxa), cuckoos (two taxa), parrots (three taxa), shorebirds (three taxa), owls (two taxa), dollarbird, kingfisher, trogon, woodpecker + toucan (two taxa), potoo, owl-nightjar, Apodiformes (two taxa), Accipitriformes (three taxa), falcons (two taxa), flamingo + grebes (three taxa), Pelecaniformes (two taxa), tubenoses (two taxa), stork, penguins (two taxa), and the loon. For rooted trees, there are 21 independent groupings, the above plus Galloanserae (five taxa). If, for example,  $k = 5$  and  $n = 40$  and the probability of observing a prespecified clade on new data where all taxa are included is

$$P_R(40, 5) = R(5) \times R(36)/R(40) \approx 3.5 \times 10^{-6}$$

or where only the independent taxonomic groups are included,

$$P_R(21, 5) = R(5) \times R(17)/R(21) \approx 6.3 \times 10^{-5}.$$

#### Predefined Clade Found as $m \geq 2$ Subgroups

The calculations can be extended to cases where the predicted clade is partitioned into  $m = 2$  or more subclades on a tree. Figure 4C and D shows two cases where a predefined clade appears in two separate areas of the tree ( $m = 2$ ). For a large number of taxa, it is still most unlikely that a pre-

defined clade will be in just two locations on a new tree. In the case shown here, there are  $k = 4$  taxa in the clade, and for  $m = 2$ , they can occur as either a single taxon and a group of three (fig. 4C) or as two groups, each with two taxa (fig. 4D). When the clade of  $k$  taxa is split into  $m$  subclades, with  $k_1, k_2, \dots, k_m$  taxa, respectively, then we must consider each combination of the  $m$  subclades separately and we find that the probability is

$$P(n, k, m) = \frac{k!B(n - k + m)}{m!B(n)} \sum_{\Sigma k_i = k} \prod_i \frac{R(k_i)}{k_i!}. \quad (5)$$

#### Literature Cited

- Barrowclough GF, Groth JG, Mertz LA. 2006. The RAG-1 exon in the avian order Caprimulgiformes: phylogeny, heterozygosity and base composition. *Mol Phyl Evol.* 41:238–248.
- Boore JL. 2006. The use of genome-level characters for phylogenetic reconstruction. *Trends Ecol Evol.* 21:439–446.
- Brown JW, Payne RB, Mindell DP. 2007. Nuclear DNA does not reconcile ‘rocks’ and ‘clocks’ in Neoaves: a comment on Ericson et al. *Biol Lett* 3:257–259.
- Brown JW, Rest JS, García-Moreno J, Sorenson MD, Mindell DP. 2008. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. *BMC Biol.* 6:6.
- Chubb AL. 2004. New nuclear evidence for the oldest divergence among neognath birds: the phylogenetic utility of ZENK (i). *Mol Phyl Evol.* 30:140–151.
- Clarke JA, Tambussi CP, Noriega JI, Erikson GM, Ketcham RA. 2004. Definitive fossil evidence for the extant avian radiation in the Cretaceous. *Nature.* 433:305–308.
- Cooper A, Penny D. 1997. Mass survival of birds across the Cretaceous-Tertiary boundary: molecular evidence. *Science.* 275:1109–1113.
- Cracraft J. 2001. Avian evolution, Gondwana biogeography and the Cretaceous-Tertiary mass extinction event. *Proc Roy Soc Lond B Biol Sci.* 268:459–469.
- Cracraft J, Barker FK, Braun M, et al. 2004. (13 co-authors). 2004. Phylogenetic relationships among modern birds (Neornithes). In: Cracraft J, Donoghue MJ, editors. *Assembling the Tree of Life*. New York: Oxford University Press. p. 468–489.
- Cracraft J, Clarke J. 2001. The basal clades of modern birds. In: Gauthier J, Gall LF, editors. *New perspectives on the origin and early evolution of birds*. New Haven (CT): Peabody Museum of Natural History, Yale University. p. 143–156.
- Curole JP, Kocher TD. 1999. Mitogenomics: digging deeper with complete mitochondrial genomes. *Trends Ecol Evol.* 14:394–398.
- de Kloet RS, de Kloet SR. 2005. The evolution of the spindlin gene in birds: sequence analysis of an intron of the spindlin W and Z gene reveals four major divisions of the Psittaciformes. *Mol Phyl Evol.* 36:706–721.
- Delsuc F, Phillips MJ, Penny D. 2003. Comment on “Hexapod origins: monophyletic or paraphyletic?” *Science.* 301:1482.
- Desjardins P, Morais R. 1990. Sequence and gene organization of the chicken mitochondrial genome. *J Mol Evol.* 32:153–161.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4:699–710.
- Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol.* 7:214.
- Dyke GJ, Cooper JH. 2000. A new psittaciform bird from the London clay (Lower Eocene) of England. *Palaeontology.* 43:271–285.

- Dyke GJ, Mayr G. 1999. Did parrots exist in the Cretaceous period? *Nature*. 399:317–318.
- Ericson PGP, Anderson CL, Britton T, Elzanowski A, Johansson US, Källersjö M, Ohlson JJ, Parsons TJ, Zuccon D, Mayr G. 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. *Biol Lett*. 2:543–547.
- Feduccia A. 2003. 'Big bang' for Tertiary birds? *Trends Ecol Evol*. 18:172–176.
- García-Moreno J, Mindell DP. 2000. Using homologous genes on opposite sex chromosomes (gametologs) in phylogenetic analysis: a case study with avian CHD. *Mol Biol Evol*. 17:1826–1832.
- Gibb GC, Kardailsky O, Kimball RT, Braun E, Penny D. 2007. Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. *Mol Biol Evol*. 24:269–280.
- Groth JG, Barrowclough GF. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. *Mol Phyl Evol*. 12:115–123.
- Guindon D, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 52:696–704.
- Hackett SJ, Kimball RT, Reddy S, et al. (17 co-authors). 2008. A phylogenomic study of birds reveals their evolutionary history. *Science*. 320:1763–1768.
- Harrison GL, McLenachan PA, Phillips MJ, Slack KE, Cooper A, Penny D. 2004. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late Cretaceous. *Mol Biol Evol*. 21:974–983.
- Harshmann J, Braun EL, Braun MJ, et al. (17 co-authors). 2008. Phylogenomic evidence for multiple losses of flight in ratite birds. *Proc Natl Acad Sci USA*. 105:13462–13467.
- Hendy MD, Penny D. 1989. A framework for quantitative study of evolutionary trees. *Syst Biol*. 38:297–309.
- Holland BR, Huber KT, Moulton V, Lockhart PJ. 2004. Using consensus networks to visualize contradictory evidence for species phylogeny. *Mol Biol Evol*. 21:1459–1461.
- Holland BR, Penny D, Hendy MD. 2003. Outgroup misplacement and phylogenetic inaccuracy under a molecular clock—a simulation study. *Syst Biol*. 52:229–238.
- Honeycutt RL, Adkins RM. 1993. Higher-level systematics of eutherian mammals—an assessment of molecular characters and phylogenetic hypotheses. *Annu Rev Ecol Syst*. 24: 279–305.
- Huelsenbeck JP, Ronquist F. 2001. MrBayes: Bayesian inference on phylogenetic trees. *Bioinformatics*. 17:754–755.
- Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol*. 23: 254–267.
- Jeffroy O, Brinkmann H, Delsuc F, Philippe H. 2006. Phylogenomics: the beginning of incongruence? *Trends Genet*. 22:225–231.
- Johansson US, Parsons TJ, Irestedt M, Ericson PGP. 2001. Clades within the 'higher land birds', evaluated by nuclear DNA sequences. *J Zool Syst Evol Res*. 39:37–51.
- Lin Y-H, McLenachan PA, Gore AR, Phillips MJ, Ota R, Hendy MD, Penny D. 2002. Four new mitochondrial genomes and the increased stability of evolutionary trees of mammals from improved taxon sampling. *Mol Biol Evol*. 19:2060–2070.
- Livezey BC, Zusi RL. 2007. Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy. II. Analysis and discussion. *Zool J Linn Soc*. 149:1–95.
- Lockhart PJ, Penny D, Hendy MD, Howe CJ, Beanland TJ, Larkum AWD. 1992. Controversy on chloroplast origins. *FEBS Lett*. 301:127–131.
- Löytynoja A, Goldman N. 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science*. 320:1632–1635.
- Matthee CA, Eick G, Willows-Munro S, Montgelard C, Pardini AT, Robinson TJ. 2007. Indel evolution of mammalian introns and the utility of non-coding nuclear markers in eutherian phylogenetics. *Mol Phylogenet Evol*. 42:827–837.
- Mayr G. 2002a. On the osteology and phylogenetic affinities of the Pseudasturidae—lower Eocene stem-group representatives of parrots (Aves, Psittaciformes). *Zool J Linn Soc*. 136: 715–729.
- Mayr G. 2002b. Osteological evidence for paraphyly of the avian order Caprimulgiforms (nightjars and allies). *J Ornithol*. 143:82–97.
- Mayr G. 2005. The postcranial osteology and phylogenetic position of the Middle Eocene *Messelastur gratulator* Peters, 1994—a morphological link between owls (Strigiformes) and falconiform birds? *J Vertebr Paleontol*. 25:635–645.
- Mayr G, Manegold A, Johansson US. 2003. Monophyletic groups within 'higher land birds'—comparison of morphological and molecular data. *J Zool Syst Evol Res*. 41: 233–248.
- Mindell DP, Sorenson MD, Dimcheff DE. 1998. Multiple independent origins of mitochondrial gene order in birds. *Proc Natl Acad Sci USA*. 95:10693–10697.
- Miyaki CY, Mitioli SR, Burke T, Wajntal A. 1998. Parrot evolution and paleogeographical events: mitochondrial DNA evidence. *Mol Biol Evol*. 15:544–551.
- Morgan-Richards M, Trewick SA, Bartosch-Härlid A, Kardailsky O, Phillips MJ, McLenachan PA, Penny D. 2008. Bird evolution: testing the Metaves clade with six new mitochondrial genomes. *BMC Evol Biol*. 8:20.
- Penny D, Foulds LR, Hendy MD. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. *Nature*. 297:197–200.
- Penny D, Hendy MD. 1986. Estimating the reliability of evolutionary trees. *Mol Biol Evol*. 3:403–417.
- Penny D, Hendy MD, Steel MA. 1991. Testing the theory of descent. In: Miyamoto MM, Cracraft J, editors. *Phylogenetic analysis of DNA sequences*. Oxford: Oxford University Press. p. 155–183.
- Penny D, Phillips MJ. 2004. The rise of birds and mammals: are microevolutionary processes sufficient for macroevolution? *Trends Ecol Evol*. 19:516–522.
- Penny D, White WT, Hendy MD, Phillips MJ. 2008. A bias in ML estimates of branch lengths in the presence of multiple signals. *Mol Biol Evol*. 25:239–242.
- Pereira SL, Baker AJ. 2006. A mitogenomic timescale for birds detects variable phylogenetic rates of molecular evolution and refutes the standard molecular clock. *Mol Biol Evol*. 23: 1731–1740.
- Philippe H, Lopez P, Brinkmann H, Budin K, Germot A, Laurent J, Moreira D, Muller M, Le Guyader H. 2000. Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proc Roy Soc Lond B Biol Sci*. 267:1213–1221.
- Phillips MJ, Delsuc F, Penny D. 2004. Genome-scale phylogeny: sampling and systematic errors are both important. *Mol Biol Evol*. 21:1455–1458.
- Phillips MJ, McLenachan PA, Down C, Gibb GC, Penny D. 2006. Combined mitochondrial and nuclear DNA sequences resolve the interrelations of the major Australasian marsupial radiations. *Syst Biol*. 55:122–137.
- Phillips MJ, Penny D. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. *Mol Phyl Evol*. 28:171–185.
- Poe S, Chubb AL. 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. *Evolution*. 58:404–415.

- Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics*. 14:817–818.
- Proakis JG, Rader CM, Ling F, Nikias CL, Moonen M, Proudler IK. 2002. *Algorithms for statistical signal processing*. Upper Saddle River (NJ): Prentice Hall.
- Rodriguez-Ezpeleta N, Brinkmann H, Roure B, Lartillot N, Lang BF, Phipippe H. 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. *Syst Biol*. 56:389–399.
- Shapiro LH, Dumbacher JP. 2001. Adenylate kinase intron 5: a new nuclear locus for avian systematics. *Auk*. 118:248–255.
- Slack KE, Delsuc F, McLenachan PA, Arnason U, Penny D. 2007. Resolving the root of the avian mitogenomic tree by breaking up long branches. *Mol Phyl Evol*. 42:1–13.
- Slack KE, Janke A, Penny D, Arnason U. 2003. Two new avian mitochondrial genomes (penguin and goose) and a summary of bird and reptile mitogenomic features. *Gene*. 302:43–52.
- Slack KE, Jones CM, Ando T, Harrison GL, Fordyce E, Arnason U, Penny D. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. *Mol Biol Evol*. 23:1144–1155.
- Sorenson MD, Oneal E, García-Moreno J, Mindell DP. 2003. More taxa, more characters: the Hoatzin problem is still unresolved. *Mol Biol Evol*. 20:1484–1499.
- Stamatakis A. 2006. RAXML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 22:2688–2690.
- Steel MA, Penny D. 1993. Distributions of tree comparison metrics—some new results. *Syst Biol*. 42:126–141.
- Stidham TA. 1998. A lower jaw from a Cretaceous parrot. *Nature*. 396:29–30.
- Susko E, Roger AJ. 2007. On reduced amino acid alphabets for phylogenetic inference. *Mol Biol Evol*. 24:2139–2150.
- Swofford DL. 2001. *PAUP\* phylogenetic analysis using parsimony (\* and other methods)*. Sunderland (MA): Sinauer Associates.
- van Tuinen M, Hedges SB. 2001. Calibration of avian molecular clocks. *Mol Biol Evol*. 18:206–213.
- van Tuinen M, Stidham TA, Hadly EA. 2006. Tempo and mode of modern bird evolution observed with large-scale taxonomic sampling. *Hist Biol*. 18:209–225.
- Waterhouse DM. 2006. Parrots in a nutshell: the fossil record of Psittaciformes (Aves). *Hist Biol*. 18:223–234.
- Whitfield JB, Lockhart PJ. 2007. Deciphering ancient rapid radiations. *Trends Ecol Evol*. 22:258–265.
- Zwickl DJ. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion [PhD dissertation]. [Austin (TX)]: The University of Texas at Austin.

Manolo Gouy, Associate Editor

Accepted October 27, 2008